

УДК 007.52

Г.И.Верченко

## ПРОБЛЕМЫ РАСПОЗНАВАНИЯ РЕЧЕВЫХ СИГНАЛОВ

ООО "Спецобъединение", г. Лиски

*В данной работе проведен анализ основных проблем, связанных с распознаванием речевых сигналов. Приведена классификация систем распознавания, указаны способы распознавания речи.*

**Ключевые слова:** распознавание, речь, язык, метод.

В течение последнего времени наблюдается развитие систем распознавания речи, которые применяются при управлении различными информационными системами. При этом используются соответствующие человеко-машинные интерфейсы.

Активные работы по распознаванию речи ведутся уже в течение нескольких десятилетий, и решение задачи весьма непростое. Это связано с тем, что даже один и тот же человек может в разных условиях произносить одно и то же слово по-разному. Не всегда представление слова в транскрипции совпадает с конкретным данным случаем произношения. Также при распознавании требуется учитывать модель языка.

Первое устройство, предназначенное для того, чтобы производить распознавание речи, было создано в 1952 году, на его основе могли распознаваться произносимые людьми цифры. Постепенно алгоритмы распознавания совершенствовались и в 90-х гг. 20 в. возникли компьютерные программы, связанные с распознаванием речи. Такие программы использовали люди, которые должны были набирать большие объемы текста. Среди указанных программ можно отметить, например, Dragon NaturallySpeaking и VoiceNavigator. Что касается надежности, то для первых версий таких программ она была не очень высока, но затем она постепенно улучшалась.

С появлением мобильных устройств, для них были созданы программы, имеющие функцию распознавания речи. Можно среди отмеченных программ указать приложение Microsoft Voice Command, позволяющее осуществлять работу с различными приложениями на основе голосового управления.

Выбор моделей распознавания речи определяется тем, какая речь - она может быть изолированной, слитной, разговорной. Качество распознавания в большой мере зависит от размеров используемого словаря, но при этом необходимо понимать, что могут появиться слова, которые близки по произношению.

Для систем распознавания речь разбивают на определенные небольшие части, которые превращаются в векторы признаков, то есть

происходит параметризация речевого сигнала. Для этих небольших частей считается, что речь является на них стационарной. Если необходимо более точное описание сигнала, то применяют процессы перекрытия.

В качестве одного из довольно распространенных подходов, связанных с распознаванием слитной (естественной) речи в настоящее время применяют скрытые марковские модели. Исследователи стремятся к тому, чтобы провести объединение скрытых марковских моделей с использованием вероятностей переходов между словами, задаваемые исходя из того какая применяется модель языка. Когда проводится распознавание естественной речи, то ориентируются на модифицированный алгоритм Витерби.

При оценке качества распознавания речи происходит сравнение последовательности распознанных слов и того, что произнес человек. В результате определяется коэффициент неверно распознанных слов.

Можно указать несколько моделей языка, которые основаны на статистическом анализе текста:

1. Модели, которые основаны на классах;
2. Триггерные модели;
3. Модели, которые основываются на частях слов;
4. Триграммная дальнедействующая модель.

Системы распознавания речи могут быть классифицированы следующим образом:

- в зависимости от размера словаря;
- зависит ли система от диктора;
- естественная (слитная) или отдельная речь;
- для каких целей предназначена создаваемая система;
- какие используются алгоритмы распознавания в системе;
- какие применяются структурные единицы;
- в зависимости от подходов, связанных с выделением структурных единиц.

Основные шаги при распознавании речи такие:

- рассмотрение качества речи;
- определение параметров в речевом сигнале, которые требуются для проведения распознавания;
- выделение фонетических и вероятностных характеристик, необходимых для того, чтобы осуществить семантический и синтаксический анализ;
- декодирование, в рамках которого происходит определение наиболее вероятной последовательности слов.

Способы распознавания речи следующие:

1. Распознавание отдельных команд;

2. Распознавание с использованием грамматики;
3. Проведение поиска ключевых слов для потока естественной (то есть, слитной) речи;
4. Проведение распознавания речи с применением большого словаря.

Перечислим некоторые основные существующие на настоящий момент дикторнезависимые системы, связанные с автоматическим распознаванием речи.

1. ACULAB - могут поддерживаться различные языки;
2. BABEAR SDK - могут поддерживаться различные языки, размер словаря может быть небольшой;
3. LOQUENDO ASR - может использоваться в телефонии;
4. LUMENVOX - дикторнезависимая система, происходит адаптация под пользователей;
5. NUANCE - высокая точность распознавания;
6. SPIRIT - поддержка различных языков, число слов может быть как фиксированным, так и произвольным

Таким образом, в данной работе проведен анализ основных особенностей, связанных с распознаванием речевых сигналов.

## ЛИТЕРАТУРА

1. Чопоров О.Н., Чупеев А.Н., Брегеда С.Ю. Методы анализа значимости показателей при классификационном и прогностическом моделировании / Вестник Воронежского государственного технического университета. 2008. Т. 4. № 9. С. 92-94.
2. Азарова Е.С. Методы фильтрации сигналов / Успехи современного естествознания. 2011. № 7. С. 64-65.
3. Свиридов В.И. Применение вейвлет-преобразований при обработке цифровых сигналов / Успехи современного естествознания. 2012. № 6. С. 110-111.
4. Гончарова Н.П. Применение методов аппроксимации данных / Успехи современного естествознания. 2011. № 7. С. 97а.
5. Головинов С.О., Миронченко С.Г., Щепилов Е.В., Преображенский А.П. Цифровая обработка сигналов / Вестник Воронежского института высоких технологий. 2009. № 4. С. 64.
6. Корчагин С.В. Разработка подсистемы анализа сигналов с помехами / Успехи современного естествознания. 2011. № 7. С. 130.
7. Завьялов Д.В. О применении информационных технологий / Современные наукоемкие технологии. 2013. № 8-1. С. 71-72.

8. Кульнева Е.Ю. Вопросы защиты информации при передаче по каналам связи / Современные наукоемкие технологии. 2013. № 8-1. С. 73.
9. Кленяева Г.В., Преображенский А.П. Современные проблемы речевой акустики и построения систем автоматического распознавания речи / Вестник Воронежского института высоких технологий. 2007. Т. 1. № 2-1. С. 71.
10. Родионова К.Ю. Использование вейвлет-преобразований для обработки цифровых сигналов / Современные наукоемкие технологии. 2013. № 8-1. С. 75.
11. Вострикова О.С. Особенности разработки подсистемы анализа сигналов с различными помехами / В мире научных открытий. 2010. № 4-3. С. 143-144.
12. Тышкевич О.В. Практическая реализация алгоритма исследования сигналов на основе методов искусственного интеллекта / В мире научных открытий. 2010. № 4-4. С. 37-39.
13. Мозговой А.А. Проблемы применения скрытых марковских моделей при распознавании рукописного текста / В мире научных открытий. 2013. № 6 (42). С. 186-198.
14. Козлова Н.А. Создание подсистемы обучения китайскому языку / Современные исследования социальных проблем. 2010. № 2. С. 93-94.
15. Преображенский Ю.П. Оценка эффективности применения системы интеллектуальной поддержки принятия решений / Вестник Воронежского института высоких технологий. 2009. № 5. С. 116-119.
16. Мозговой А.А. Распознавание рукописных текстовых символов, вводимых в мобильные устройства / Вестник Воронежского института высоких технологий. 2011. № 8. С. 48-50.
17. Преображенский Ю.П., Ермаченко А.С. Сравнительный анализ алгоритмов поиска текстовых фрагментов / Вестник Воронежского института высоких технологий. 2010. № 7. С. 76-78.
18. Тышкевич О.В., Преображенский А.П. Реализация алгоритма исследования сигналов на основе методов искусственного интеллекта / Вестник Воронежского института высоких технологий. 2009. № 5. С. 199-204.
19. Федоринова Н.В., Преображенский А.П. Исследование исправляющей способности кодов / Вестник Воронежского института высоких технологий. 2008. № 3. С. 094-097.
20. Ермолова В.В., Преображенский Ю.П. Методика построения семантической объектной модели / Вестник Воронежского института высоких технологий. 2012. № 9. С. 87-90.

21. Зяблов Е.Л., Преображенский Ю.П. Марковские процессы принятия решений первого типа с несколькими поглощающими состояниями / Вестник Воронежского института высоких технологий. 2010. № 6. С. 68-71.
22. Зазулин А.В., Преображенский Ю.П. Особенности построения семантических моделей предметной области Вестник Воронежского института высоких технологий. 2008. № 3. С. 026-028. /
23. Преображенский Ю.П. Разработка методов формализации задач на основе семантической модели предметной области / Вестник Воронежского института высоких технологий. 2008. № 3. С. 075-077.
24. Кульнева И.Ю., Гащенко И.А. Характеристики вейвлет-преобразования при решении задач обработки сигналов / Вестник Воронежского института высоких технологий. 2013. № 10. С. 149-152.
25. Преображенский А.П. Анализ методов кодирования разных видов информации / Вестник Воронежского института высоких технологий. 2013. № 10. С. 74-77.
26. Свиридов В.И. Основные характеристики методов распознавания голоса / Моделирование, оптимизация и информационные технологии. 2013. № 2. С. 4.
27. <http://www.speechpro.ru/technologies/recognition>
28. <http://www.slideshare.net/VictorOsetrov/2012-22967611>
29. Benesty J., Sondhi M., Huang Y. (eds.) Springer Handbook of Speech Processing. Springer, 2008, 1176 p.
30. Лобанов Б.М., Цирульник Л.И. Компьютерный синтез и клонирование речи. Минск: Белорусская наука, 2008, 344 с.

G.I. Verchenko

## THE PROBLEMS OF RECOGNITION OF SPEECH SIGNALS

*Joint-stock company «Spetsobedineniye», Liski*

*In this paper, the analysis of the main issues associated with recognition of speech signals is given. The classification of systems of recognition, methods of speech recognition is presented.*

**Keywords:** recognition, speech, language, method.